OPTIMAL MARGINAL TAX RATES FOR LOW INCOMES: POSITIVE, NEGATIVE, OR ZERO?

BY STEFAN HOMBURG

DISCUSSION PAPER NO. 255

MAY 2002

ISSN 0949 – 9962

ABSTRACT: Previous studies have shown that the optimal marginal tax rate at the bottom of the income distribution may be positive, negative, or even zero. This paper reexamines this problem in a unified framework and tries to evaluate the arguments. It turns out that the case for positive marginal tax rates is quite strong, whereas negative marginal tax rates seem to be based on a misconception.

KEYWORDS: Optimal income taxation, contract theory, incentives, welfare.

JEL-CLASSIFICATION: H2, D3

1. INTRODUCTION

Designing income support programs is still an important matter of concern. The theory of nonlinear income taxation, initiated by Mirrlees (1971), approaches this problem as follows: It considers an abstract tax schedule, embracing both positive tax liabilities in the ordinary sense as well as transfer payments which are conceived of as negative taxes. The analysis aims at characterizing the qualitative properties of such income tax schedules. Since negative taxes are transfer payments, it yields recommendations regarding the optimal design of income support programs as a by-product.

However, these recommendations are anything but clear-cut. They can be categorized into three groups. The first group, which includes Mirrlees' original paper, states that optimal marginal income tax rates at the bottom of the income distribution should be *strictly positive*. In fact, Mirrlees has proven that marginal income tax rates are always non-negative. But his simulations produced marginal tax rates which are strictly positive at the bottom and decreasing on their entire domain, contradicting the accustomed ideal of tax rate graduation. Optimal marginal tax rates at zero income varied between 21% and 50% in his study. These results agree with the high implicit tax rates of many contemporary welfare programs.

The second group, including Seade (1977, 1982), argues that the optimal marginal tax rate should be *zero* at the bottom of the income distribution in order to leave labor supplies of the least skilled persons undistorted. Since the same result holds undisputedly at the top of the income distribution, Seade suggested that optimal tax schedules typically take an "S-form", with marginal tax rates increasing from zero to some maximum and then decreasing again to zero. His no-distortion-at-the-bottom result harmonizes with the basic exemptions integrated in most income tax schedules.

The third group contends that marginal tax rates at the bottom may well be *strictly negative*. These results go back to Diamond (1980) who used the Mirrlees model, too, but made one important change in the assumptions: Instead of considering persons that can adjust their labor supplies smoothly, Diamond assumed fixed working hours, implying that people are confronted with binary choices only. Negative marginal tax rates characterize the well-known U. S. earned income tax credit (EITC).

Thus, the results obtained so far exhaust the entire range of possibilities and leave the policymaker helpless. The objective of this paper is to discuss these inconsistent outcomes in a unified framework and to evaluate them. The framework used is a finite variant of the standard model which assumes an uncountable infinity of taxpayers. With infinitely many taxpayers, economic reasoning is liable to become replaced by purely mathematical operations which do not tell anything about the economic forces at work. By contrast, the finite model allows deriving all results referred to above in a simple and lucid fashion. Moreover, some limit characterizations will be provided which clarify the relationship between the optimal tax formulae of the standard model and their finite counterparts.

Section 2 describes the basic model. Section 3 reports on a number of established theorems. The main sections 4 to 6 are devoted to replicating the findings of the positive, zero, and negative marginal tax rates, respectively. Section 7 concludes.

2. THE MODEL

There are finitely many persons who have different skills and thus earn different wage rates per hour, w^h, where $h = 0, 1 ... H$ and $0 \leq w^0 < w^1 < ... < w^H$. In order to derive limit properties later on, it is convenient to assume a *limit* probability distribution function F which possesses a continuous density $f > 0$ on its support $[w^0, w^H]$. The actual skill distribution of the finite economy is given by probability masses $f^0 = F(0) > 0$ and $f^h = F(w^h) - F(w^{h-1})$ for all $h > 0$. Each person with wage rate w^h (person h, for short) chooses some *consumption* c^h and some *labor supply* ℓ^h such that the *commodity bundle* (c^h, ℓ^h) belongs to the uniform consumption space $\mathcal{C} = \mathcal{R}_{o+} \times [0; \ell^{max}]$, where $\ell^{max} > 0$. A person's *gross income* is denoted as $y^h = w^h \ell^h$.

The uniform utility function u: $\mathcal{C} \rightarrow \mathcal{R}$ is continuous, strictly monotonically increasing in c, strictly monotonically decreasing in ℓ and strictly concave on its entire domain. At least in the interior, it is twice continuously differentiable with partial derivatives $u_c(c, \ell) > 0$ and $u_\ell(c, \ell) < 0$ and a negative definite Hessian. Moreover, consumption is a gross substitute for leisure (meaning that in the absence of taxes, consumption is an increasing function of the wage rate), and leisure is non-inferior.

An *allocation* is a vector $(c^h, \ell^h)_{h=0...H}$ in $\mathcal{C}^{H+1}$. The *social objective,* which can be interpreted as an expected utility representation of a man choosing the optimal tax schedule behind a veil of ignorance, reads

$$(1) \qquad \underset{(c^h, \ell^h)_{h=0...H} \in \mathcal{C}^{H+1}}{max!} \quad \sum_{h=0}^{H} u(c^h, \ell^h) f^h .$$

Aggregate (per capita) consumption is $\Sigma c^h f^h$. Assuming a linear technology, aggregate (per capita) output is $\Sigma y^h f^h$. With an exogenous *tax revenue* $g > 0$, also defined in per capita terms, an allocation must satisfy the *resource constraint*

$$(2) \qquad \sum_{h=0}^{H} (y^h - c^h) f^h \geq g .$$

As the differences $y^h - c^h = T(y^h)$ are in fact tax payments, this inequality represents the government's budget constraint. A *tax schedule* is a mapping $\mathcal{R}_{o+} \rightarrow \mathcal{R}$ which associates a tax payment T(y) with every income and thus confronts the persons with legal choices (c, y). Since wage rates are exogenous, the preference ordering over commodity bundles (c, ℓ) induces a preference ordering over *pairs* (c, y) of consumption and income which depends on the respective wage rate. Some person k accepts the pair (c^k, y^k) intended for him only if no other pair (c^h, y^h) exists which he prefers strictly. Otherwise person k will *mimic* person h by choosing c^h instead of c^k and y^h/w^k instead of ℓ^k. Hence, any *feasible allocation* must satisfy the resource constraint and the *self-selection constraints*

$$(3) \qquad u(c^k, \ell^k) \geq u(c^h, y^h/w^k) \text{ for all k and h where } y^h/w^k \leq \ell^{max}.$$

A self-selection constraint is called *downward* if $k > h$ and *upward* if $k < h$; it is called *adjacent* if $k = h \pm 1$. A *second-best optimum* maximizes the social objective (1) subject to the resource constraint (2) and the self-selection constraints (3).

This model is identical to the standard model introduced by Mirrlees (1971), except that there are finitely many tax payers. The standard model can be obtained as a limiting case in the following way. With w^0 and w^H fixed and H increasing, consider an equidistant partition of the interval $[w^0, w^H]$, i. e. a finite set of wage rates $w^0 < w^1 < ... < w^H$ such that $w^h - w^{h-1} = \delta > 0$ for all $h > 0$. The economy is completely described by the partition, the limit distribution function, the utility function, and the required tax revenue. A *sequence of increasingly fine economies* is a sequence with the property $\delta \to 0$. As δ converges to zero, the differences between adjacent skill levels become smaller and smaller, and so do the probability masses f^h, whereas the ratios $f^h/\delta = [F(w^h) - F(w^{h-1})]/(w^h - w^{h-1})$ converge to the density f for all $h > 0$. Thus, the limit distribution becomes approximated by an increasing sequence of step functions.

3. PRELIMINARY RESULTS

The following lemma simplifies the problem considerably since the bulk of the self-selection constraints is eliminated and the remaining constraints are mostly equalities rather than weak inequalities.

LEMMA 1: The original maximization problem (1) to (3) is equivalent to the transformed maximization problem:

$$\tag{4} \begin{aligned} &\underset{(c^h,\ell^h)_{h=0\ldots H}\in\mathcal{C}^{H+1}}{max!} \quad \sum_{h=0}^{H} u(c^h,\ell^h)\, f^h \\ &s.t. \quad i) \quad \sum_{h=0}^{H} (y^h - c^h)\, f^h = g, \\ &\qquad\quad ii) \quad u(c^h,\ell^h) = u(c^{h-1}, y^{h-1}/w^h) \quad \text{for all } h > 0, \\ &\qquad\quad iii) \quad y^h \geq y^{h-1} \quad \text{for all } h > 0. \end{aligned}$$

This lemma, proven in the appendix, is easy to understand. It says, firstly, that a second-best optimum is always production efficient in the sense that the resource constraint holds with equality. Secondly, all adjacent downward self-selection constraints are binding because the optimization mechanism redistributes as much as possible from top to bottom. At the optimum, every person $h > 0$ is indifferent between the pair (c^h, y^h) intended for him and the pair (c^{h-1}, y^{h-1}) intended for his left-hand neighbor. This is referred to as the *chain property* of second-best allocations. Finally, income is monotonically increasing in the wage rate. Since the final inequalities are weak, it may occur that persons with distinct wage rates have the same income. This is usually referred to as *bunching*.

In the absence of distortionary taxes, each person's marginal rate of substitution $-u_\ell/u_c$ equals the wage rate w at any interior solution. A positive marginal tax rate implies that the marginal rate of substitution falls short of the wage rate and vice versa. Therefore *implicit marginal tax rates* can be defined as follows:

$$\tag{5} L^h = 1 + \frac{u_\ell(c^h, \ell^h)}{u_c(c^h, \ell^h)w^h} \quad \text{and} \quad R^h = 1 + \frac{u_\ell(c^h, y^h/w^{h+1})}{u_c(c^h, y^h/w^{h+1})w^{h+1}}.$$

The variable L^h denotes the implicit marginal tax rate of person h, whereas R^h denotes the implicit marginal tax rate of person h+1 if the latter has the same income as the former. This occurs when person h+1 mimics person h or when there is bunching at income y^h. The two implicit marginal tax rates play the central role in our model. There is an important relationship between them. To derive it, substitute $\ell = y/w$ in the utility function and select some $\bar{u}$ from the utility function's range. The equation $u(c, y/w) = \bar{u}$ defines an implicit function c(y) with the derivative

$$\text{(6)} \qquad c'(y) = -\frac{u_\ell(c(y), y/w)}{u_c(c(y), y/w)\, w} > 0 .$$

This derivative represents the marginal rate of substitution in c-y-space. From definitions (5) it is clear that the implicit marginal tax rates L^h and R^h equal $1 - c'(y^h)$, evaluated at w^h and w^{h+1}, respectively. The mean value theorem implies existence of a wage rate w between w^h and w^{h+1} such that

$$\text{(7)} \qquad R^h - L^h = \sigma^h (w^{h+1} - w^h) > 0 \quad \text{where} \quad \sigma^h := -\frac{dc'(y^h)}{dw} > 0 .$$

The positive sign of σ^h follows from the gross substitutability assumption. Thus, indifference curves in c-y-space become flatter if the wage rate increases, see figure 1. Considering a tax schedule $T(y) = y - c(y)$ with slope $T'(y) = 1 - c'(y)$ it follows at once that the schedule becomes steeper at a certain income level if the wage rate increases. This is a result of fundamental importance: R^h will always exceed L^h at any positive income level; at zero income the two are identical by definition.

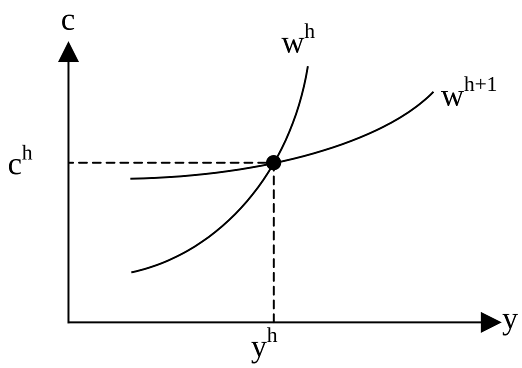

FIGURE 1 – Indifference Curves in y-c-space.

LEMMA 2: Any second-best optimum can be supported by a continuous tax schedule which has left-derivatives L^h and right-derivatives R^h at all incomes $y^h > 0$ $(h < H)$.

An example of such a tax schedule for an economy with four persons is depicted in figure 2. The kinks at income levels y^1 and y^2 are points of non-differentiability. For instance, the left-derivative at income y^1 equals L^1 and the right-derivative equals R^1, and as $R^1 > L^1$, the kink is upward. The tax schedule is differentiable at zero income because here, only a right-derivative exists. A formal proof of lemma 2 can be found in Homburg (2002).

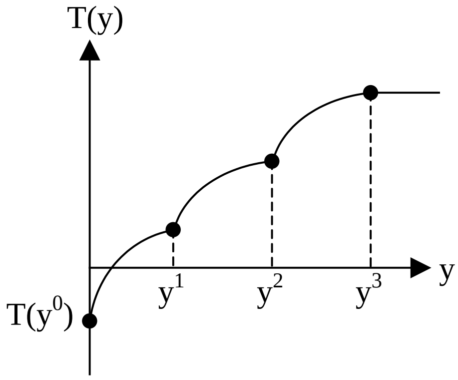

FIGURE 2 – Optimal Tax Schedule.

Before concluding this section we would like to relate the implicit marginal tax rates to the concept of the *deadweight loss* which plays such an important role in the theory of commodity taxation. Assume that person $h < H$ has strictly positive consumption and income and faces a marginal tax rate $L^h = 0$ at the outset. Introducing a small distortion means reducing consumption and leisure in such a way as to leave the resource constraint unaffected. Thus, consider a number $\varepsilon > 0$ and the associated utility $U(\varepsilon) = u(c^h - \varepsilon, \ell^h - \varepsilon/w^h)$ and differentiate with respect to ε to obtain the marginal reduction in utility,

(8) $$U'(0) = -u_c(c^h, \ell^h)\, L^h .$$

With a lump-sum tax, utility falls by $-u_c(c^h, \ell^h)$. Therefore, $L^h = u_c L^h/u_c$ represents the marginal deadweight loss, expressed as a mark-up. As usual, the deadweight loss vanishes if $L^h = 0$. The variable R^h has an analogous interpretation. Let $\hat{U}(\varepsilon) = u(c^h - \varepsilon, (y^h - \varepsilon)/w^{h+1})$ denote the utility of person h+1 if he mimics person h. Differentiating again yields

(9) $$\hat{U}'(0) = -u_c(c^h, y^h/w^{h+1})\, R^h .$$

Since a lump-sum tax diminishes utility by $-u_c(c^h, y^h/w^{h+1})$, R^h is the marginal deadweight loss imposed on person h+1 if the latter mimics person h, expressed again as a mark-up. From inequality (7) one infers that a distortion imposed on person h harms the potential mimicker h+1 more than it harms person h himself. The intuition behind this result is obvious: If person h+1 accepted the new pair $(c^h - \varepsilon, (y^h - \varepsilon)/w^{h+1})$, he would suffer from the same decrease in consumption as person h but would only gain ε/w^{h+1} extra units of leisure, whereas person h gains ε/w^h extra units.

For lack of a better term, $R^h - L^h$ is called the *differential deadweight loss*. To reiterate, the differential deadweight loss is strictly positive at any strictly positive income, indicating that a distortion at income y^h harms person h+1 more than it harms person h. In particular, the differential deadweight loss is strictly positive if L^h itself vanishes, i.e. if person h's choice is not distorted at the margin. In the following section it will become clear that the differential deadweight loss is the driving force behind the optimality of positive marginal tax rates in the finite model, but not so in the continuum model.

4. POSITIVE MARGINAL RATES

To avoid technical problems which have already been dealt with in the literature, we assume that there is no bunching at the optimum, i.e. $y^h > y^{h-1}$ for all $h > 0$, and that individual optima are interior. Consequently, the final constraints in the transformed optimization problem (4) drop out and all remaining constraints are equalities. Moreover, it is easy to show that the resource constraint and the adjacent self-selection constraints are linearly independent. This allows using a conventional Lagrangean approach without apology, since the familiar first-order conditions are necessarily fulfilled at an optimum, irrespective of the problem's convexity properties (Bertsekas 1999, proposition 3.1.1), and thus can be used to characterize the solutions. Introducing shadow prices λ and μ^h for the resource constraint and the adjacent self-selection constraints, respectively, the Lagrangean function reads:

$$(10) \qquad \mathcal{L} = \sum_{h=0}^{H} u(c^h, \ell^h) f^h + \lambda \sum_{h=0}^{H} \left(y^h - c^h - g\right) f^h + \sum_{h=1}^{H} \mu^h \left(u(c^h, \ell^h) - u(c^{h-1}, y^{h-1}/w^h)\right).$$

All shadow prices are strictly positive at an optimum because any slack in a constraint would allow increasing the social objective. Differentiating with respect to c^h and ℓ^h,

$$(11) \qquad \begin{aligned} \frac{\partial \mathcal{L}}{\partial c^h} &= u_c(c^h, \ell^h) f^h - \lambda f^h + \mu^h u_c(c^h, \ell^h) - \mu^{h+1} u_c(c^h, y^h/w^{h+1}) = 0, \\ \frac{\partial \mathcal{L}}{\partial \ell^h} &= u_\ell(c^h, \ell^h) f^h + \lambda w^h f^h + \mu^h u_\ell(c^h, \ell^h) - \mu^{h+1} u_\ell(c^h, y^h/w^{h+1}) \frac{w^h}{w^{h+1}} = 0, \end{aligned}$$

and rearranging terms yields

$$(12) \qquad (f^h + \mu^h)\, u_c(c^h, \ell^h) = \mu^{h+1}\, u_c(c^h, y^h/w^{h+1}) + \lambda f^h,$$

$$(13) \qquad (f^h + \mu^h) \frac{u_\ell(c^h, \ell^h)}{w^h} = \mu^{h+1} \frac{u_\ell(c^h, y^h/w^{h+1})}{w^{h+1}} - \lambda f^h.$$

For $h = H$, the terms including μ^{h+1} vanish, and dividing the two equations gives $L^H = 0$. This is the familiar result that no distortion should be imposed at the very top of the income distribution. For every person $h < H$, adding the two equations and substituting L^h and R^h from (5) yields

$$(14) \qquad (f^h + \mu^h)\, u_c(c^h, \ell^h)\, L^h = \mu^{h+1} u_c(c^h, y^h/w^{h+1})\, R^h.$$

This equation is interesting because it exhibits three impacts of distortionary taxes: Firstly, taxing person h reduces individual utility by $u_c L^h$ and aggregate utility by $f^h u_c L^h$. This is an *efficiency cost*. Secondly, the reduction in h's utility tightens the self-selection constraint which prevents h from mimicking h–1. Therefore, $\mu^h u_c L^h$ represents a *redistributive cost*. Finally, slack emerges in the self-selection constraint which prevents h+1 from mimicking h because the tax lowers the potential mimicker's utility, too. Multiplying this slack by the shadow price of the subsequent self-selection constraint, one obtains the *redistributive gain* $\mu^{h+1} u_c R^h$. An optimal income tax schedule must balance the efficiency and redistributive cost on the one hand and the redistributive gain on the other.

Equation (14) implies that L^h and R^h have the same sign but does not help to determine it. To do so, use (12) to substitute $\mu^{h+1} u_c(c^h, y^h/w^{h+1})$ by $(f^h + \mu^h) u_c(c^h, \ell^h) - \lambda f^h$, subtract and add $\lambda f^h L^h$ on the left-hand side of (14) and solve for L^h to obtain

(15) $$L^h = \frac{(f^h + \mu^h) u_c(c^h, \ell^h) - \lambda f^h}{\lambda f^h} (R^h - L^h) > 0 .$$

The sign is strictly positive because the numerator equals $\mu^{h+1} u_c > 0$ and $R^h - L^h$ has already been shown to be strictly positive. Thus, the optimal marginal rate is the product of two effects:

– A *redistribution effect* which equals the ratio on the right-hand side. This ratio represents the social rate of return of increasing person h's consumption. Such an increase costs λf^h, raises aggregate utility by $f^h u_c$ and allows additional redistribution toward persons below h, which is worth $\mu^h u_c$.
– An *incentive effect* $R^h - L^h$, corresponding to the differential deadweight loss defined in the preceding section. Imposing a distortion on person h harms the potential mimicker h+1 more than it harms person h himself and therefore enables additional redistribution.

The total effect is the product of these partial effects. It is strictly positive since the government wishes to redistribute from top to bottom and a positive distortion enables it to do so.

Equation (15), which is the principal result for the finite economy, states that the left-derivative of an optimal tax schedule must be strictly positive at every income $y^h > 0$, except at the highest income where it vanishes. Since $R^h > L^h$, the same is true for the right-derivatives, and if all incomes happen to be strictly positive, the characterization is complete. However, it may occur that y^0 vanishes, in which case the above derivation does not apply. As bunching was assumed absent, y^1 will be strictly positive, and from the chain property one knows that person 1 is indifferent between the pairs (c^1, y^1) and $(c^0, 0)$. Moving along the indifference curve of person 1 from the former pair to the latter, the marginal rate of substitution increases, implying $R^0 > L^1$. Therefore, if the smallest income happens to vanish, the marginal tax rate at the bottom of the income distribution equals $T'(0) = R^0 > L^1 > 0$, where the equality is clear from figure 2 and the last inequality follows from (15). In this case the tax schedule is even differentiable at the bottom. If, on the other hand, the smallest income is strictly positive, which presumes that the smallest wage rate is strictly positive, then the tax schedule is non-differentiable at the bottom, but the left- and right-derivatives are both strictly positive.

In the finite economy, the sign of the marginal tax rate equals the sign of the differential deadweight loss under any redistributive motive. It is a remarkable fact that this incentive effect goes astray as the economies become increasingly fine: If w^{h+1} approaches w^h, the right-derivative R^h approaches the left-derivative L^h, as is clear from definitions (5). However, this does not imply that the marginal tax rates converge to zero since formula (15) shows that the redistribution effect diverges: As f^h becomes arbitrarily small in the limit, increasing consumption of some person $h > 0$ costs almost nothing, but doing so has a finite value even in the limit because it enables redistributing more to people below h. In order to determine the tax rate's limit behavior, use (7) to express $R^h - L^h$ as $\sigma^h(w^{h+1} - w^h) = \sigma^h \delta$, substitute into (15) and put δ into the denominator of the redistribution effect:

(16)
$$L^h = \frac{(f^h + \mu^h)u_c(c^h, \ell^h) - \lambda f^h}{\lambda f^h/\delta}\sigma^h \to \frac{\mu^h u_c \sigma^h}{\lambda f}.$$

The limit expression is identical with Mirrlees' (1971) equation (27) or Seade's (1977) equation (19), which were derived using the maximum principle and partial integration. The present derivation makes use of the fact that f^h converges to zero whereas f^h/δ converges to the density f when the economies become increasingly fine. It corroborates Seade's intuition that the marginal tax rate has only a redistributive function and not an incentive function in the continuum model. Hence the economic stories behind the positive marginal rate in the finite model on the one hand and in the continuum model on the other are quite different.

5. Zero Marginal Rates

The main result of the preceding section confirms the previous findings of Stiglitz (1982) and Guesnerie and Seade (1982) who showed that the marginal tax rate at the bottom of the income distribution will always be strictly positive in the finite model, provided that the tax system wishes to redistribute toward the lower skilled. At the same time it runs counter to Seade's (1982) conclusion that the marginal tax rate vanishes at the bottom of the income distribution under certain assumptions. Essentially, c^0 and y^0 are strictly positive in his model (implying that w^0 is strictly positive), and bunching is excluded. Under these premises, equation (15) holds for person 0 but the shadow price μ^0 vanishes: Preventing person 0 from mimicking his left-hand neighbor has no social value because the latter does not exist. Substituting this and $R^h - L^h = \sigma^h \delta$ into (15) and eliminating f^h yields

(17)
$$L^0 = \frac{u_c(c^0, \ell^0) - \lambda}{\lambda}\sigma^0\delta \to 0 \ ,$$

if all variables except δ remain finite in the limit, as Seade assumed. At the bottom of the income distribution, the incentive effect converges to zero again but the redistribution effect remains finite, giving rise to a vanishing combined effect. In the preceding section it was shown that the redistribution effect diverges at any skill level exceeding w^0 as it enables additional redistribution to persons *below* h at virtually no cost. At skill level w^0, where there are no such persons, the argument becomes void, and thus undistortionary taxation is optimal.

In the continuum model, the "no distortion at the bottom and at the top" results follow from the transversality conditions which state that the Lagrangean multiplier μ(w) vanishes at both ends of the income distribution. In the finite model, these transversality condition read $\mu^0 = 0$ and $\mu^{H+1} = 0$, respectively. As persons numbered "–1" and H+1 do not exist, these multipliers are not present in the Lagrangean (10), implying that the corresponding marginal tax rates vanish.

Despite the seeming symmetry of the arguments, the "no distortion at the top" result is relatively robust whereas its counterpart, regarding the marginal tax rate at the bottom, is not. In deriving the latter one must assume that the persons with the lowest skill level have a strictly positive income at the optimum. This presupposes $w^0 > 0$ but the latter assumption is not sufficient: At the bottom of the optimal income distribution, there will normally be an interval of nonworkers even if all skill levels were strictly positive. Accordingly, all simulation studies –

including Mirrlees (1971), Stern (1976), Tuomala (1990), or Kanbur and Tuomala (1994) – arrive at strictly positive tax rates at the bottom. Moreover, one must keep in mind that (14) is a local result. Even if it holds, marginal tax rates will be strictly positive in any right-hand neighborhood of the smallest income.

6. NEGATIVE MARGINAL RATES

The standard model of optimal income taxation presumes that people can vary their effort continuously over some interval $[0, \ell^{max}]$. This is referred to as an *intensive choice*. Diamond (1980) argued that at least for some people the only relevant choice may be whether or not to work at all (*extensive choice*). The following analysis takes up this suggestion and assumes $\ell^h \in \{0, \overline{\ell}\}$ for all h, where $\overline{\ell}$ is a strictly positive number. Perhaps in order to keep his continuum model differentiable, Diamond also assumed that workers differ in the disutility of work. As the finite model does not rely on differentiability assumptions, such a premise is not required here and will not be made: We would like to study in isolation the effects resulting from substituting extensive for intensive choices.

Diamond's model of extensive choices makes it necessary to assume $w^0 = 0$, since mimicking would otherwise present no problem and a first-best tax schedule could be implemented. It is important to keep in mind that with intensive choices excluded, a person with wage rate $w^h > 0$ can only pretend to be totally unable to earn positive income, thus choosing $\ell^h = 0$ and mimicking person 0; he cannot mimic other persons with distinct positive wage rates. As a result, the self-selection constraints (3) must be replaced by the conditions

$$u(c^k, \ell^k) \geq u(c^h, 0) \text{ for all k and h.} \tag{18}$$

Finding a second-best tax schedule means maximizing (1) subject to (2) and (18) under the assumptions $w^0 = 0$ and $\ell^h \in \{0, \overline{\ell}\}$ for all h, but with all other assumptions introduced in section 2 unchanged. Every second-best tax schedule has the following three features:

Firstly, all nonworkers enjoy the same consumption. This follows directly from the preceding self-selection constraints: If the government offered two pairs with distinct consumption quantities and zero income, everybody would choose the pair with the higher consumption. As person 0 has zero income, all nonworkers receive c^0.

Secondly, all workers enjoy the same consumption, say $\overline{c}$. For, assume there exist two workers k and h such that the first has a higher consumption at the optimum. Since an optimum must satisfy the self-selection constraints (18), we have $u(c^k, \overline{\ell}) > u(c^h, \overline{\ell}) \geq u(c^0, 0)$. Redistributing some consumption from k to h in accordance with the resource constraint will not violate the self-selections constraints and will certainly increase the social objective because person k had a lower marginal utility of consumption at the outset.

Thirdly, workers are indifferent between working and becoming unemployed; the respective self-selection constraints hold as equalities:

$$u(\overline{c}, \overline{\ell}) = u(c^0, 0) \ . \tag{19}$$

This is easy to see: At any feasible allocation, workers' consumption $\overline{c}$ must exceed nonworkers' consumption c^0 since otherwise everybody would prefer to be unemployed. As a consequence, workers have a lower marginal utility of consumption than nonworkers. This is obvious in the case of an additive utility function but also true if leisure is non-inferior (Homburg 2002, lemma 2). Hence, redistributing consumption from workers to nonworkers increases the social objective, and the government will do so as much as possible, equation (19) characterizing the state where no more redistribution is possible without violating the self-selection constraints. To summarize, a second-best optimum is characterized by two consumption levels only, one for the workers and one for the nonworkers, and by the remarkable fact that all persons enjoy the same utility.

As people cannot vary their choices locally, the implicit marginal tax rates cease to be meaningful, but optima can still be characterized using *discrete marginal tax rates*

$$m^h = \frac{T(y^h) - T(y^{h-1})}{y^h - y^{h-1}}, \tag{20}$$

which are defined for all $y^h > y^{h-1}$. Because workers' net incomes are all identical whereas gross incomes differ according to the different wage rates, it follows immediately that for any two adjacent persons h and $h-1$ both working, the discrete marginal tax rate m^h equals one hundred per cent. For any two adjacent persons both not working, the marginal rate is not defined. The interesting marginal tax rate at the bottom of the income distribution is encountered where $y^h > 0$ and $y^{h-1} = 0$.

PROPOSITION: At any second-best optimum, the discrete marginal tax rate m^h is non-negative for $y^h > 0$ and $y^{h-1} = 0$.

PROOF: Since h is a worker and h–1 is a nonworker, $T(y^h) = y^h - \overline{c}$ and $T(y^{h-1}) = -c^0$. Thus, if the marginal rate $m^h = (y^h - \overline{c} + c^0)/y^h$ is strictly negative at the alleged optimum, $\overline{c} - c^0$ exceeds y^h. Switching to $\ell^h = 0$ and $c^h = c^0$ yields an output surplus because the fall in aggregate consumption exceeds the fall in aggregate output. This move is feasible and does not change person h's utility because of (19). The emerging surplus can be used to make all persons better off, contradicting the optimality of the original allocation. Hence, $m^h \geq 0$. *Q.E.D.*

At the optimum, nonworkers are indexed 0...h–1 and workers are indexed h...H. The separating index $h > 0$ always exists since $w^0 = 0$ implies existence of at least one zero income, whereas $g > 0$ implies existence of at least one strictly positive income. The proposition says that it does not pay to "bribe" a nonworker to join the labor force if the required increase in consumption exceeds the resulting increase in output. In the non-generic case, $m^h = 0$ and it is immaterial whether or not person h works; the changes in aggregate consumption and aggregate output just balance. Generically, m^h will be strictly positive. Therefore, the optimality of non-negative marginal tax rates still holds if labor supply decisions are extensive rather than intensive. This is at variance with Diamond's (1980) result.

To contrast Diamond's findings with the above proposition, assume a Cobb-Douglas utility function $u(c, \ell) = [c(500-\ell)]^{0.4}$, fixed working hours $\overline{\ell} = 250$, and a per capita tax revenue $g = 100$. The tables display five hourly wage rates distributed uniformly; f^h equals 20 per cent for all h.

Table I shows the optimal distribution of consumption (net income) which was calculated subject to a given distribution of gross income. Utility $u = 194.5$ is the same for all persons. The marginal tax rate equals –6 per cent at the bottom and one hundred per cent otherwise.

Compared with the least skilled person, person 1 receives an extra consumption amounting to $2111-1056=1055$. The bottom marginal tax rate is *strictly negative* because this extra consumption exceeds his gross income. Diamond, optimizing only with respect to consumption in his theoretical derivation, presented a similar example.

TABLE I
NEGATIVE MARGINAL TAX RATE

w	c	ℓ	y	T	m
0	1056	0	0	-1056	--
4	2111	250	1000	-1111	-6%
8	2111	250	2000	-111	100%
12	2111	250	3000	889	100%
16	2111	250	4000	1889	100%

The important message of the proposition above is that negative marginal tax rate can never be optimal. Table I does not describe the full optimum. The full optimum, resulting from optimizing over consumption *and* labor supply, is shown in table II, where person 1 no longer joins the labor force. All utilities have risen from 194.5 to 195.0, implying that this move represents a Pareto-improvement. The economic explanation is obvious: If person 1 becomes unemployed, an output surplus emerges which can be used to make all persons (including person 1 himself) better off. Every rational taxpayer or welfare recipient would prefer the allocation depicted in table II to the negative marginal tax case depicted in table I.

TABLE II
OPTIMAL INCOME TAX SCHEDULE

w	c	ℓ	y	T	m
0	1062	0	0	-1062	--
4	1062	0	0	-1062	--
8	2125	250	2000	-125	47%
12	2125	250	3000	875	100%
16	2125	250	4000	1875	100%

Introducing a negative marginal tax rate such as the U. S. earned income tax credit (EITC) means moving from table II to table I. While this move makes all persons worse off, output and employment increase. This suggests that in evaluating such a reform measure one must carefully distinguish *welfare effects* on the one hand and *output and employment effects* on the other (Browning, 1995). Most empirical analyses of the EITC (like Hotz and Scholz, 2001) concentrate solely on the latter, which gives rise to a bias in favor of the EITC. Keynes once coined the famous phrase that one can always ensure full employment by letting the unemployed dig something into the earth and dig it out thereafter, paying them an income for this useless activity. To do so would certainly not be a sensible policy.

Using a variant of Diamond's (1980) model, Saez (1999) defended the optimality of negative marginal tax rates with the following argument:

"Starting from a situation with equal transfers to low income workers and the unemployed, increasing the transfers to the low income workers is beneficial from a pure redistributive point of view ... This also encourages some of the unemployed to join the labor force at zero fiscal cost as transfers are initially equal for the two groups. As a result, it is unambiguously welfare enhancing to provide a larger transfer to low income workers than to the unemployed." (Saez 1999, p. 12).

This is clearly wrong. Assume transfers to persons 0 and 1 are equal at the outset. Introducing a strictly negative marginal tax rate means offering person 1 an extra consumption, relative to the unemployed person 0, which exceeds the extra output resulting from his effort. If person 1 accepts and starts working, this induces a strictly positive fiscal cost. The additional cost necessitates reducing other persons' utilities. As (19) implies that all utilities are identical at the optimum, a Pareto deterioration results. From the taxpayers' perspective, it is better to accept and to finance a certain degree of voluntary unemployment than to fight this unemployment using negative marginal tax rates. To be sure, this reasoning does not reject negative *taxes* (wage subsidies) in general; it only rejects negative *marginal tax rates*. From table II, which describes the full optimum, one infers that the person with the smallest positive income receives a wage subsidy amounting to 125. The marginal tax rate equals 47 per cent at the bottom.

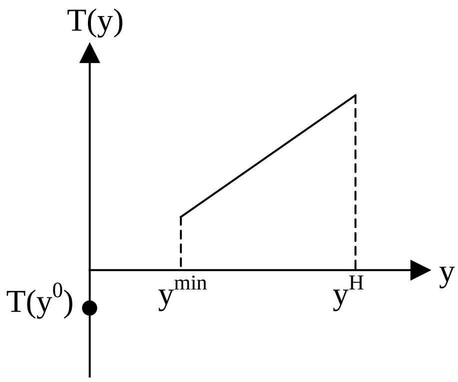

FIGURE 3 – Limit Tax Schedule.

When the economies become increasingly fine, the tax schedule converges to a limit such as depicted in figure 3. The limit schedule consists of an isolated point representing the transfer to the unemployed and a straight line, starting at some minimum observable income y^{min}. The tax $T(y^{min})$ may be positive or negative, but must exceed $T(y^0)$. The slope of the line equals unity. Marginal tax rates of one hundred per cent are an unrealistic feature, of course, triggered by the assumption that people cannot find part time jobs or otherwise reduce their efforts, which itself is unrealistic.

7. CONCLUSION

This paper examined the sign of optimal marginal tax rates at the bottom of the income distribution. The framework used was the finite counterpart of the standard model of the optimal income tax. Our results sustain the view that the marginal income tax rate should be strictly positive at the bottom because it has an incentive function in the finite model and at least a redistributive function in the continuum model. Zero marginal rates are more or less theoretical artifacts: In the finite model, the marginal rate will never vanish at the bottom, whereas in the continuum model it will do so only under stringent assumptions, and just for a null set of persons.

Considering extensive instead of intensive labor supply choices does not change these results: Zero marginal rates are still non-generic, and strictly negative rates are never optimal. These conclusions have been derived, of course, within the usual framework and may be upset once different assumptions become introduced. However, the premises of the usual framework are quite general and convincing.

Negative marginal tax rates such as the EITC seem to stem from a multilateral *illusion*, where the poor believe they profit from the subsidy while the rich are happy about the seeming relief from taxation which they think will accompany the fall in unemployment. Thus there is unanimous political support, and as long as economists endorse this illusion, the EITC will have a bright future. But the analysis conducted in section 6 showed that inducing low-skilled people to join the labor force by means of negative marginal tax rates results in a Pareto deterioration. From this perspective, taxpayers are better advised to accept a certain degree of voluntary unemployment.

Faculty of Business Administration and Economics, University of Hannover, Königsworther Platz 1, D-30167 Hannover, Germany. www.fiwi.uni-hannover.de.

APPENDIX

PROOF OF LEMMA 1: Propositions 2a), 4, and 1b) in Homburg (2002) show that every solution of the original maximization problem must be production efficient and must satisfy the chain property and the monotonicity property. Hence, the constraints i) to iii) of the transformed optimization problem are indeed necessary for an optimum. It remains to be shown that all allocations satisfying these constraints belong to the feasible set defined by (2) and (3).

To do so we make use of Homburg's (2002) *agent monotonicity*: If a person is indifferent between two distinct pairs $(\overline{c}, \overline{y})$ and $(\underline{c}, \underline{y})$, where the former contains more income than the latter, then every person with a higher wage rate strictly prefers $(\overline{c}, \overline{y})$ and every person with a lower wage rate (that can reach both pairs) strictly prefers $(\underline{c}, \underline{y})$. This follows from the gross substitutability assumption in the text.

The binding self-selection constraint $u(c^h, \ell^h) = u(c^{h-1}, y^{h-1}/w^h)$ states that person h is indifferent between (c^h, y^h) and (c^{h-1}, y^{h-1}). For all numbers $i = 1...h-1$, the preceding self-selection constraints assert that person $h-i$ is indifferent between (c^{h-i}, y^{h-i}) and (c^{h-i-1}, y^{h-i-1}). Now, agent monotonicity implies that person h prefers every (c^{h-i}, y^{h-i}) to every (c^{h-i-1}, y^{h-i-1}) because compared to person $h-i$ he has a higher wage rate and because the former pairs contain more income than the latter; the preference is even strict if any two pairs under consideration are distinct. From the transitivity of the preference ordering it follows that person h prefers the pair intended for him to all pairs intended for low-skilled persons. Therefore, all downward self-selection constraints are satisfied. As a perfectly analogous argument holds for the upward self-selection constraints, every allocation satisfying i) to iii) belongs to the feasible set. i.e. satisfies all constraints of the original problem. Q.E.D.

REFERENCES

Bertsekas, Dimitri P. (1999): *Nonlinear Programming*. Belmont: Athena Scientific.

Browning, E. K. (1995) Effects on the Earned Income Tax Credit on Income and Welfare. *National Tax Journal*, 14, 23–43.

Diamond, P. A. (1980): “Income Taxation with Fixed Hours of Work”, *Journal of Public Economics,* 13, 101–110.

Guesnerie, R. and J. Seade (1982): “Nonlinear Pricing in a Finite Economy”, *Journal of Public Economics,* 17, 157–179.

Homburg, St. (2002): “The Optimal Income Tax: Restatement and Extensions”, Discussion Paper, University of Hannover.

Hotz, V. J. and J. K. Scholz (2001): “The Earned Income Tax Credit.” NBER Working Paper 8078.

Mirrlees, J. A. (1971): “An Exploration in the Theory of Optimum Income Taxation”, *Review of Economic Studies*, 38, 175–208.

Kanbur, R. and M. Tuomala (1994): “Inherent Inequality and the Optimal Graduation of Marginal Tax Rates”, *Scandinavian Journal of Economics,* 96, 275–282.

Mirrlees, J. A. (1971): “An Exploration in the Theory of Optimum Income Taxation”, Review of Economic Studies, 38, 175–208.

Saez, E. (1999): “Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses”, NBER Working Paper 7708.

Seade, J. K. (1977) “On the Shape of Optimal Tax Schedules”, *Journal of Public Economics*, 7, 203-235.

Seade, J. K. (1982): “On the Sign of the Optimum Marginal Income Tax”, *Review of Economic Studies*, 49, 637–643.

Stern, N. (1976): “On the Specification of Models of Optimum Income Taxation”, *Journal of Public Economics,* 6, 123–162.

Stiglitz, J. E. (1982): “Self-selection and Pareto Efficient Taxation”, *Journal of Public Economics,* 17, 213-240.

Tuomala, M. (1984): “On the Optimal Income Taxation”, *Journal of Public Economics*, 23, 351–366.